Hannes Leitgeb
What theories of Truth should be like

Por Ariel Roffé

Estructura de la presentación:

1) Ocho requisitos deseables para una teoría de la verdad

2) Combinaciones maximales posibles de requisitos

3) Conclusiones

4) Extras

1) Ocho requisitos

¿Cómo *debería ser* una teoría de la verdad?

¿Cuáles son los requisitos que, *prima facie*, querríamos que una teoría de la verdad cumpla?

Clarificación previa: ¿A qué nos referimos cuando hablamos de una "teoría de la verdad"?

Hay dos tipos de teorías de la verdad: filosóficas y lógicas.

Leitgeb se va a ocupar de las segundas. Podría decirse que las teorías lógicas de la verdad se ocupan solo de caracterizar el comportamiento inferencial del predicado de verdad, cuáles son sus características "estructurales" o cómo podemos razonar con él.

Una teoría lógica de la verdad no se ocuparía de la "naturaleza" de la verdad, de especificar tipo de relación metafísica que hay entre una oración verdadera y el mundo –si es que alguna–, etc.

Sin embargo, no hay que tomarse muy en serio esa distinción ya que una teoría lógica de la verdad tiene que ser defendida mediante argumentos filosóficos. La adecuación de una teoría de la verdad no es un asunto puramente formal.

(a) La verdad debería ser expresada por un predicado

A la verdad se la puede tratar formalmente por medio de un operador o un predicado.

Un operador se aplica a oraciones. Por ejemplo, podríamos introducir como un operador de verdad a la doble negación (satisface el esq-T: $\neg\neg A \equiv A$)

Un predicado se aplica a términos (constantes de individuo o términos de función) que denotan miembros del dominio. Puesto que la verdad se dice de oraciones, tendremos que tener en el dominio *códigos* (o nombres) de oraciones, a fin de poder aplicarles el predicado de verdad.

Ventajas de usar un predicado: posibilidades expresivas más amplias.

- Podemos cuantificar adentro de posiciones predicativas, lo cual permite decir cosas como "todos los teoremas de la aritmética son verdaderos":

$\forall x\ (\ Bew(x) \supset Tr(x)\ )$, expresando propiedades metateóricas (corrección) adentro del lenguaje

- Podemos tener términos de función adentro de un predicado, lo cual permite decir cosas como "la última oración dicha por César es verdadera" sin saber cuál es esa oración.

(a) La verdad debería ser expresada por un predicado

Puesto que el predicado de verdad se aplica a oraciones (es mejor esto que aplicarlo a entidades filosóficas problemáticas como las proposiciones), entonces debe haber una teoría de la sintaxis clara que nos diga cuáles son las fórmulas bien formadas y las oraciones del lenguaje.

Para lenguajes de primer orden esto se cumple.

(b) Si el predicado de verdad se añade a una teoría T, debemos poder probar todos los teoremas de T como verdaderos

Formalmente: $\vdash_T A \Rightarrow \vdash_{T + Tr} Tr(\text{'}A\text{'})$

Este es un requisito bastante obvio. Si PA prueba "2 + 2 = 4" y nuestra teoría de la verdad no nos dice que "Tr('2 + 2 = 4')" entonces algo anda mal. Podemos considerar a nuestra teoría como incompleta (no prueba algo que es verdadero).

Llamemos a esto "completitud débil".

** Podríamos agregar como complemento a otro requisito de "completitud fuerte":

(b') $M \vDash_T A \Rightarrow \vdash_{T + Tr} Tr(\text{'}A\text{'})$ donde M es el modelo pretendido o estándar de la teoría T

Esto es distinto al requisito débil, porque los teoremas de Gödel nos muestran que para cualquier teoría lo suficientemente expresiva:

$M \vDash_T A$ no implica que $\vdash_T A$

Es decir, hay "oraciones indecidibles", que son verdaderas en el modelo estándar pero que *no se pueden probar*. O sea, que hay "verdades indemostrables".

(b) Si el predicado de verdad se añade a una teoría T, debemos poder probar todos los teoremas de T como verdaderos

** También podríamos agregar los correspondientes requisitos de corrección.

(b'') "corrección débil":

$\vdash_{T+Tr}$ Tr('A') $\Rightarrow$ M$\vDash_T$ A   Si Tr('A') es teorema de T + Tr entonces A es verdadera y en el modelo estándar de T

$\vdash_{T+Tr}$ ¬Tr('A') $\Rightarrow$ M$\nvDash_T$ A

(b''') "corrección fuerte":

$\vdash_{T+Tr}$ Tr('A') $\Rightarrow$ $\vdash_T$ A   y

$\vdash_{T+Tr}$ ¬Tr('A') $\Rightarrow$ $\nvdash_T$ A

Nótese que los teoremas de Gödel recién mencionados implican que una teoría de la verdad que satisface completitud fuerte no puede satisfacer corrección fuerte (el primer requisito en particular).

Otro requisito emparentado es el de (b'''') conservatividad:

$\vdash_{T+Tr}$ A $\Rightarrow$ $\vdash_T$ A (Agregar un predicado de verdad no permite probar cosas nuevas del vocabulario de base)

(c) El predicado de verdad no debe estar tipeado

En principio parece un paso menor pasar de que "2 + 2 = 4 es verdadero" a que "es verdadero que '2 + 2 = 4 es verdadero' ". Parece que, *prima facie*, no hay ninguna razón para pensar que el predicado de verdad debe estar tipeado

Es más, Kripke da buenas razones para pensar que el predicado de verdad del lenguaje natural no está tipeado

Por ejemplo, si lo estuviera, no podríamos decir cosas como "todo lo que dijo Eduardo es verdadero" sin conocer lo que dijo y los índices que usó.

También tenemos casos como el de Nixon en los que una persona no podría hablar acerca de todo lo que dijo la otra (no hay una asignación de índices tal que ambos puedan decir lo que quieren decir).

El motivo de Tarski para postular esto fue la ocurrencia de paradojas semánticas. Pero acá estamos preocupados solo por lo que querríamos en principio, olvidándonos de los problemas que puedan traer o no.

(d) Todos los bicondicionales-T deben ser derivables

La idea de Tarski era que la derivabilidad de todos los bicondicionales T a partir de una definición del concepto de verdad era condición necesaria para la asignación de la extensión correcta al predicado de verdad.

Otros (Revisión) creen que el esquema-T *es la definición* del concepto de verdad.

Si uno cree que A y Tr(‘A’) deben ser intersubstituibles en todo contexto no oblicuo, y la lógica que tiene valida “si A entonces A”, los bicondicionales T también se siguen.

En general la derivabilidad de los bicondicionales T es aceptado como un requisito. En lo que no hay acuerdo es en qué significa el “si y solo si” (para algunos es una equivalencia material, para otros una equivalencia definicional o una conectiva no clásica)

(d) Todos los bicondicionales-T deben ser derivables

** Nuevamente, parece que hay otros requisitos más débiles que podrían resultar suficientes para los propósitos filosóficos de este.

(d') Por ejemplo, en lugar de que las instancias del esquema T salgan como teoremas, podríamos pedir que valgan las siguientes reglas de inferencia:

$$\frac{A}{Tr(\text{'}A\text{'})} \qquad \frac{Tr(\text{'}A\text{'})}{A}$$

(llamemos a esto transparencia fuerte)

(d'') Podríamos exigir que valieran solo las siguientes reglas (transparencia débil):

$$\frac{\vdash_{T+Tr} A}{\vdash_{T+Tr} Tr(\text{'}A\text{'})} \qquad \frac{\vdash_{T+Tr} Tr(\text{'}A\text{'})}{\vdash_{T+Tr} A}$$

(puede ocurrir que algún modelo base –y por lo tanto de T+Tr– satisfaga algo como G y no satisfaga Tr('G'))

(e) La verdad debería ser composicional

Que una propiedad X sea composicional quiere decir que:

La posesión de X por parte de una oración molecular depende únicamente de la posesión de X por parte de sus oraciones componentes y el modo como estas se componen.

Por ejemplo, que el significado sea composicional, nos explica la productividad del lenguaje (que seamos capaces de comprender el significado y de formular oraciones nuevas, que nunca habíamos escuchado y que nadie nunca antes había formulado).

Que la verdad sea composicional significa que el valor de verdad de una oración depende del valor de verdad de sus oraciones componentes.

Queremos que la verdad de Tr('A') dependa del valor de verdad de A. Así, si el valor de verdad de A es falso, queremos que ¬Tr('A') sea verdadera, y viceversa. Es decir, queremos que valga algo como esto: $\forall x\, (\, Sent_L(x) \supset (\, Tr(\text{'}\neg x\text{'}) \equiv \neg Tr(\text{'}x\text{'})\,)\,)$

Los axiomas de T(PA) consistían en bicondicionales T para oraciones atómicas del lenguaje base + clausulas de composicionalidad como esta para el resto de las conectivas y los cuantificadores + el esquema-T irrestricto.

(f) La interpretación pretendida no debería estar excluida como modelo de la teoría de la verdad.

Sin una interpretación, una oración de un lenguaje formal no es más que una serie de símbolos sin ningún significado.

Para que, bajo una interpretación, una oración formal como “$\forall x\ Tr(x)$” signifique “Todas las oraciones son verdaderas” hay ciertas cosas que no deben ocurrir:

Por ejemplo, no debe ocurrir que en la extensión de Tr(x) haya cosas que no son códigos de oraciones –por ejemplo, números no estándar–.

Pero vimos que hay teorías como FS tales que *todos* sus modelos tienen algún número no estándar en la interpretación de Tr.

Por lo tanto, la interpretación pretendida (aquella según la cual “Tr” significa “verdad”) no es un modelo de FS. Esto parece altamente indeseable, ya que parece querer decir que FS no puede hablar acerca de la verdad.

(f) La interpretación pretendida no debería estar excluida como modelo de la teoría de la verdad.

Todas las teorías tienen modelos que no son el pretendido (lo máximo que se puede expresar es hasta el isomorfismo, en 2° orden). Si no tenemos alguna otra forma aceptable de especificar cuál de todas esas es la interpretación correcta, podríamos decir que la referencia de los términos de la teoría "ambigua", que la Tr no se refiere solo a la verdad, sino que podríamos estar hablando acerca de otras cosas.

Casos como el de FS son peores porque el predicado "Tr" no puede referirse en absoluto a la verdad. El problema no es que hay otras interpretaciones que hacen a los axiomas de la teoría verdaderos, sino que la que la única que verdaderamente nos importa no los hace.

(g) La “lógica interna” y la “lógica externa” deberían coincidir

La “lógica externa” refiere al conjunto de las leyes lógicas que valen “afuera” del alcance de un predicado de verdad.

La “lógica interna” al conjunto de las leyes lógicas que valen “adentro” de tales.

Hay teorías en las que ambas cosas no coinciden.

Por ejemplo (axiomatizaciones de los modelos de Kripke con K3?), prueban $A \vee \neg A$ para toda oración A, pero no prueban Tr(‘$A \vee \neg A$’) para algunas instancias de A (por ejemplo λ).

En casos como ese, la lógica externa es clásica, pero la interna no.

Parece que decir de una oración que es verdadera de alguna manera cambia la lógica que la rige, lo cual es indeseable.

(g) La “lógica interna” y la “lógica externa” deberían coincidir

Este requisito parece seguirse del de completitud.

Si las leyes lógicas de la teoría T son parte de T, entonces que una de ellas no sea probada como verdadera en T + Tr haría fallar completitud.

Sin embargo, según el autor, este requisito agrega algo más:

“(g) ensures that this is the case for *all* logical truths whatsoever; if truth is to be explained for a language with truth predicate –which consequently would not be a language of a purely mathematical or empirical theory but would be partly semantic– then the same logical truths involving ‘Tr’ should show up in the outer and the inner logic.”

**Hay dos interpretaciones posibles de esto:

(g_1) Si se toma a T + Tr como una “lógica de la verdad” entonces tiene que probar como verdaderos a sus propios teoremas (tiene que valer d’’).

(g_2) Si se toma a T + Tr no como parte de la lógica, sino como una teoría, entonces lo que T + Tr tiene que validar son las leyes de (su?) lógica subyacente instanciadas con el predicado Tr. (si A v ¬A es una ley lógica, T+Tr tiene que validar Tr(‘B’) v ¬Tr(‘B’))

(g) La “lógica interna” y la “lógica externa” deberían coincidir

** (g_2) me parece la interpretación más adecuada de lo que está diciendo.

El problema sería –lo que tiene que validar como verdaderas son las leyes de su propia lógica subyacente instanciadas con Tr, o las leyes lógicas de T instanciadas con Tr.

La respuesta a esta pregunta es que *ambas*. Leitgeb está trabajando desde un marco de teoría de la prueba. Una teoría axiomática de la verdad se obtiene agregando un predicado de verdad a otra teoría T, gobernado por ciertos axiomas.

Por lo tanto, puesto que tiene como teoría subyacente a T, tiene como lógica subyacente a la lógica subyacente de T.

Lo que podría resultar extraño es que los modelos de T + Tr pueden no ser clásicos (al menos en el modo en que se interpreta al predicado Tr). Sin embargo las reglas de inferencia que valen en el sistema axiomático (que se usan para derivar teoremas a partir de los axiomas) sí son las clásicas.

(h) La lógica externa debe ser clásica

Este me parece el requisito más debatible de todos. Leitgeb da varios argumentos en su favor:

1) - Si una teoría es revisada, lo es en base a otra teoría que se mantiene fija (que actúa como teoría subyacente suya).

- La lógica es teoría subyacente de toda otra teoría, si se modifica la lógica no hay nada más que pueda mantenerse fijo.

- La lógica clásica es la opción "por default", aquella que es usada más ampliamente en filosofía, ciencias empíricas, etc.

** Nuevamente creo que se ve que Leitgeb trabaja en teoría de la prueba. Es cierto que los sistemas axiomáticos se construyen unos sobre otros (y no hasta el infinito).

Si uno lo ve desde la perspectiva de la teoría de modelos, esto no es cierto. No hay una teoría que actúe como fundamento de todas las demás (para caracterizar a las verdades lógicas se utiliza teoría de conjuntos, y para razonar en teoría de conjuntos se usa la lógica). Más que un fundacionismo hay "circularidad".

(h) La lógica externa debe ser clásica

** Uno podría objetarle al autor que:

- Nociones de teoría de la prueba como "derivabilidad" también utilizan vocabulario conjuntístico ("secuencia de oracionés …"), por lo que no hay nada absolutamente fundamental, que no presuponga otra cosa.

- No es cierto que siempre las teorías se revisan en base a otras teorías que se mantienen fijas. Hay mucha gente que propone revisar la lógica por "motivos intrínsecos" a la lógica (por ejemplo, creen que la lógica clásica valida ciertas cosas que no deberían validarse)

- Su argumento, de corte Quineano ("nada es inmodificable, pero tocar la lógica hace que se caigan muchas más cosas que tocar alguna otra teoría"), presupone algún tipo de monismo lógico (que no pueden haber revisiones locales de la lógica, sino que toda revisión es global)

** Lo que rescato de este argumento es que: si uno compra que la lógica por default es la clásica, que el *único* motivo para revisar la lógica son paradojas que surgen en una teoría de la verdad, y hay otra teoría de la verdad que no cambia la lógica pero que lidia igualmente bien con ellas, ha de preferirse esta última.

2) Combinaciones maximales posibles

No se pueden satisfacer todos los requisitos anteriores a la vez.

Tarski probó que una teoría con (a) + (b) + (c) + (d) + (h) es inconsistente.

Si seguimos el principio Quineano de "mutilación mínima" tenemos que preguntarnos cuál de las cosas que podemos abandonar nos haría perder menos, esto es, cuál es la combinación que satisface la máxima cantidad de requisitos posibles

La solución de Tarski consistió en abandonar (c), satisfaciendo todo el resto de los requisitos: (a) + (b) + (d) + (e) + (f) + (g) + (h) TOTAL: 7

Sin embargo, la cuestión no parece ser solamente cuántos requisitos podemos satisfacer a la vez, ya que distintos requisitos parecen tener distinta importancia relativa. Importa también cuáles abandonamos.

Según Leitgeb, (c) es un requisito muy importante (la solución de Tarski no es la mejor). También lo son (a) y (b). Con lo cual, deberíamos ver qué pasa si abandonamos (d) o (h).

2.1) Abandonar (d) y sostener (h)

Si abandonamos (d), ¿Cuántos requisitos pueden mantenerse?

¿Los otros siete juntos?

NO. McGee (1985) probó que (a) + (b) + (c) + (e) + (g) + (h) hace fallar (f).

Quedan entonces tres opciones:

Opción 1)	Abandonar (f):	(a) + (b) + (c) + (e) + (g) + (h)
Opción 2)	Abandonar (g):	(a) + (b) + (c) + (e) + (f) + (h)
Opción 3)	Abandonar (e):	(a) + (b) + (c) + (f) + (g) + (h)

Las tres opciones son realizables. Ejemplos:

De 1): Axiomatización del revisionismo FS

De 2): Axiomatización de Kripke con un esquema de valuación K3 (Kleene Fuerte)

De 3): Axiomatización de Kripke con un esquema supervaluacionista

Las tres opciones satisfacen 6 de los 8 requisitos.

2.2) Abandonar (h) y sostener (d)

Si abandonamos (h), ¿Cuántos requisitos pueden mantenerse?

¿Los otros siete juntos?

Quizás sí. La teoría de Field es un buen candidato a satisfacer los otros siete, aunque está aun está abierto a discusión si satisface composicionalidad con respecto al condicional que introduce (si satisface (e))

3) Conclusiones

** La sensación que queda después de lo que acaba de exponer, es que Leitgeb cree que la teoría de Field (si es que logra cumplir sus promesas) es la mejor alternativa.

Sin embargo, no afirma esto sino que sus conclusiones son más modestas. No pretende haber resuelto la cuestión sino solo dado un marco para pensarla. Ese marco podría complejizarse, por ejemplo:

- Poniendo algún ranking de importancia de requisitos (o incluso, asignar un valor numérico de importancia).

- Podrían considerarse grados en los que una teoría no satisface alguno.

- Podrían agregarse requisitos adicionales (por ejemplo, pragmáticos), o revisar/quitar/modificar algunos de los que él propuso

4) Extras

** Hay algunas combinaciones de requisitos más (entre los que yo presenté y que Leitgeb no examina) que no son posibles:

- Una teoría puede aspirar a satisfacer como máximo o bien completitud fuerte y corrección débil ((b') y (b'')), o bien completitud débil y corrección fuerte ((b) y (b''')). El primer tipo de teorías va a ser más fuerte (va a probar más cosas) que el segundo, y a mi juicio es preferible. Por ejemplo:

    - Cuando hacemos una teoría de la verdad aritmética queremos probar las cosas que son verdades aritméticas. Las verdades aritméticas son aquellas que son verdaderas en el modelo pretendido.

    - Si hay oraciones (sobre todo del lenguaje base) tales que ni ellas ni sus negaciones son declaradas como verdaderas, entonces hay un sentido importante en el que la teoría de la verdad es incompleta (no hay oraciones del lenguaje base de la aritmética que tengan un valor de verdad indeterminado)

Esto también es prácticamente realizable. Tanto FS como T(PA) son fuertemente completas y débilmente correctas.

4) Extras

** Sin embargo, una teoría que satisface completitud fuerte y corrección débil, no puede satisfacer a la vez Transparencia (fuerte o débil –(d') o (d'')) y conservatividad (b'''').

¿Por qué no? Al menos para teorías de primer orden podemos razonar así:

Los teoremas de Gödel implican que toda teoría consistente lo suficientemente expresiva tiene oraciones indecidibles (como G, tal que no se prueba ni G ni ¬G).

Pero los modelos clásicos son bivalentes. Eso quiere decir que en todo modelo oraciones como G son o bien verdaderas o bien falsas (en algunos es verdadera, en otros es falsa). En particular, en el modelo estándar debe ser o bien verdadera, o bien falsa (de hecho, es verdadera).

Pero si es verdadera en el modelo estándar, por completitud fuerte, T + Tr debe probar Tr('G'). Y puesto que vale transparencia (débil o fuerte) debe probar G. Puesto que G era indecidible para la aritmética, T + Tr no es conservativa.

Puesto que transparencia es altamente deseable, tenemos que elegir entre:

-Completitud fuerte + Corrección débil + No conservatividad y

-Completitud débil + Corrección fuerte + Conservatividad

4) Extras

[FIN]